PATENT APPLICATION

PRODUCTION OF RECOMBINANT LACTOFERRIN AND LACTOFERRIN POLYPEPTIDES USING cDNA SEQUENCES IN VARIOUS ORGANISMS

This invention was made with government support under Grant No. HD27965 awarded by the National Institute of Health. The government has certain rights in the invention.

RELATED APPLICATIONS

This application is a continuation in part of pending application serial no. 07/967,947, filed 10/27/92, which in turn is a continuation of application serial no. 07/348,270, filed 05/05/89, now abandoned. This application is also a continuation in part of pending application serial no. 07/873,304 filed 04/24/92.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates generally to the field of iron-binding glycoproteins. More specifically, the present invention relates to the recombinant production of various lactoferrins.

2. Description of the Prior Art

Lactoferrin (LF) is an iron-binding glycoprotein found in milk and other secretions and body fluids. It is one of a number of iron binding proteins, sometimes referred to as transferrins, and is involved in iron binding and delivery in mammals.

Human lactoferrin (hLF) is a member of the transferrin family of iron-binding monomeric glycoproteins. It was originally discovered in milk where it can reach levels of 7 grams/liter in colostrum. LF has since been detected in other external fluids of humans and other mammals. The fluids include tears, saliva and mucosal secretions and also in the secondary granules of polymorphonuclear leukocytes.

Lactoferrin has been implicated as a factor in resistance against enteritis infections in suckled newborn humans. The bacteriocidal/bacteriostatic actions are considered to be due at least in part to the iron binding properties of lactoferrin. Lactoferrin decreases the iron availability to iron-requiring microorganisms and thereby interferes with their growth and reproduction. At least one non-ironbinding bactericidal domain has also been reported for human lactoferrin. Lactoferrin is also considered to have antiviral properties and to have other potential therapeutic applications.

LF is a 78 kilo Dalton (k Da) glycoprotein having a bilobal structure with a high degree of homology between the C and N terminal halves which is evident at both the amino acid and three dimensional structural level. Each of these lobes can reversibly bind one ferric iron with high affinity and with the concomitant binding of bicarbonate. The biological functions proposed for lactoferrin include protection against microbial infection, enhanced intestinal iron absorption in infants, promotion of cell growth, regulation of myelopoiesis and modulation of inflammatory responses.

Human lactoferrin (hLF) has a high affinity for iron and two Fe^{3+} cations can be bound per molecule. The complete HLF protein has been subjected to amino acid sequencing and is reported to have 703 amino acids. There are two glycosylation sites. Metz-Boutigue *et al.*, *Eur. J. Biochem.*, **145**:659-676 (1984). Anderson *et al.*, *Proc. Nat'l Acad. Sci.* USA, **84**:1769-1773 (April 1987).

In other studies, a cloned cDNA probe for amino acids 428 to 703 of the Metz-Boutigue structure of the lactoferrin protein was isolated. The cDNA sequence was in general agreement with the earlier analysis of the amino acid sequence of the protein. Rado *et al.*, *Blood*, **79**; **4**:989-993, **79**; **4**:989-993 (Oct. 1987). The probe was reported to encompass approximately 40% of the coding region and the 3' terminus. The cDNA sequence for both porcine, Lydon, J. P., *et al.*, *Biochem. Biophysic. ACTA*, **1132**:97-99 (1992); Alexander, L. J., *et al.*, *Animal Genetics*, **23**:251-256 (1992) and bovine lactoferrin, Mead, P. E., *et al.*,

Nucleic Acids Research, **18**:7167 (1990); Pierce, A., *et al.*, *Eur. J. Biochem.*, **196**:177-184 (1991), have been determined.

Polypeptides derived from lactoferrin are also known to be biologically active. A fragment containing a possible iron binding site was reported by Rado, et al. *supra*. An N-terminal human lactoferrin fragment, including a bactericidal domain of HLF, was isolated from a pepsin digest. Bellamy, W.M., *et al.*, *Biochem. Biophys. ACTA*, **1121**:130-136 (1992). Synthetic 23 and 25 amino acid polypeptides were synthesized and found to have activities similar to the fragments derived by pepsin digestion. The synthesis details, yields and purity of the synthetic peptides were not reported. Bellamy *et al.* do not provide a practical route to large scale production of the polypeptides free of the contaminates resulting form isolation from natural products.

The bactericidal domain from lactoferrin has a broad spectrum of antimicrobial action. Bellamy, W.M. *et al.*, *J. App. Bact.* **73**, 472-479 (1992). Although Bellamy *et al.* report that bovine lactoferrin isolated from milk can provide commercial quantities of the bovine polypeptide by pepsin digestion, the materials used in both studies had a minimum purity of only 95%. Bellamy, *et al.* do not provide constructs for the large scale production of synthetic human or bovine lactoferrin or lactoferrin polypeptides. Neither does Bellamy *et al.* provide the ability to produce peptides that are not available by enzyme digestion.

Filamentous fungi have been successfully employed as hosts in the industrial production of extracellular glycoproteins. Certain industrial strains are capable of secreting gram quantities of these proteins. In addition, filamentous fungi are able to correctly perform post-translational modifications of eucaryotic proteins and many strains have U.S. Food and Drug Administration approval. Furthermore, large scale fermentation technology and downstream processing experience is available.

Currently, there is no efficient and economical way to produce hLF, other species lactoferrin, or to control production of lactoferrin polypeptides. Consequently, a long felt need and description in this art would be met by the

development of an efficient method for the production of human lactoferrin for nutritional and therapeutic applications and for further investigation into its mechanism of action.

SUMMARY OF THE INVENTION

The invention comprises the verified cDNA sequences for human lactoferrin, and cDNA expression systems for use of various lactoferrin DNA sequences to produce human, bovine, porcine and other lactoferrins for a variety of end uses. The cDNA expression systems of the invention also provide a practical route and method to make lactoferrin polypeptides or fragments having biological activity. The hLF cDNA includes an open reading frame of 2133 nucleotides coding for a protein of 711 amino acids. These 711 amino acids include 19 amino acids corresponding to a secretion signal peptide sequence followed by 692 amino acids of mature human lactoferrin. The cDNA sequence and deduced amino acid sequence differ from the previously published data of Metz-Boutigue, *supra*.

In one embodiment, the present invention provides for a recombinant plasmid comprising the cDNA of human or other lactoferrin. The plasmid of the present invention is adapted for expression in a eucaryotic cell and contains the regulatory elements necessary for the expression of the human lactoferrin cDNA in this eucaryotic cell.

In another embodiment, the present invention provides for a transformed cell which includes a heterologous DNA sequence which codes for lactoferrin or a polypeptide related to lactoferrin. The heterologous DNA sequence will preferably be incorporated into a plasmid. Eucaryotic host cells are selected from the group consisting of mammalian cells, immortalized mammalian cells, fungi or yeasts. Preferred cells include filamentous fungi comprising *Aspergillus*, and yeasts. The plasmid contains a plasmid vector into which a polydeoxyribonucleotide (DNA) segment coding for human or other lactoferrin protein has been inserted.

In yet another embodiment of the present invention, there is provided a process for producing recombinant human or other lactoferrin which comprises culturing a transformant eucaryotic cell, which includes a recombinant plasmid. The plasmid contains a plasmid vector having a polydeoxyribonucleotide coding for the lactoferrin protein. After culturing in a suitable nutrient medium until lactoferrin protein is formed, the lactoferrin protein is isolated.

In still yet another embodiment of the present invention, there is provided a recombinant expression vector. This vector comprises a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression; (2) cDNA coding for lactoferrin; (3) appropriate transcription and translation initiation and termination sequences; and (4) a genetic element for selection of transformed cells or spores such as *Aspergillus* spores that have been transformed with the vector.

In still yet another embodiment of the present invention, there is provided a method for producing biologically active recombinant lactoferrin. The method comprises synthesizing sequences containing a selectable marker gene, a promotor, a transcription termination sequence, and a linker sequence; cloning the sequences to form a plasmid; digesting the plasmid with a restriction endonuclease; inserting a cDNA coding for lactoferrin into a restriction site; and transforming eucaryotic cells with the plasmid expressing lactoferrin cDNA.

BRIEF DESCRIPTION OF THE DRAWINGS

So that the manner in which the above recited features, advantages, and objects of the invention, as well as others which will become clear, are obtained and can be understood in detail, more particular descriptions of the invention briefly summarized above may be had by reference to certain embodiments thereof which are illustrated in the appended drawings. These drawings form a part of this specification.

It is to be noted, however, that the appended drawings illustrate preferred embodiments of the invention and therefore not to be considered limiting

of its scope. The invention may admit to other equally effective equivalent embodiments.

Fig. 1 is a schematic drawing of the hLF cDNA including the locations of the 5' untranslated region, the secretion peptide signal sequence, mature lactoferrin and 3' untranslated region.

Fig. 2 is the cDNA sequence (SEQ. ID No. 1) with deduced amino acids (SEQ. ID No. 2) for the human lactoferrin protein and signal peptide sequence.

Fig. 3 is a schematic representation of an autoradiograph of recombinant human lactoferrin protein expressed from the complete cDNA.

Fig. 4 is a schematic representation of an autoradiograph of the results of *in vitro* translation of a 2,140 bp human lactoferrin sequence and hLF protein in reticulocyte lysates.

Fig. 5 depicts a schematic representation of the *Aspergillus oryzae* expression plasmid, pAhlfg.

Fig. 6 shows a southern blot analysis of transformed *Aspergillus oryzae* strains.

Fig. 7 depicts an RNA analysis of transformant versus control A07.

Fig. 8 shows the silver stained SDS-acrylimide gel analysis of recombinant LF secretion and purification.

Fig. 9 illustrates the characterization of recombinant human LF.

Fig. 10 is a western immunoblot of cellular extracts of transformed *E.coli* cells expressing the C terminal fragment of LF.

Fig. 11 shows the coomassie-stained SDS-PAGE analysis of extracts of transformed *E.coli* cells expressing the C terminal fragment of LF.

Fig. 12 shows the expression and purification of the glutathione S-transferase/LFN-1 fusion protein.

Fig. 13 Schematic representation of the *A. Oryzae* universal expression plasmid, pAG.

Fig. 14 is the (A) cDNA sequence (SEQ. ID No. 3) with (B) deduced amino acids (SEQ. ID No. 4) for the bovine lactoferrin protein.

Fig. 15 is the (A) cDNA sequence (SEQ. ID No. 5) with (B) deduced amino acids (SEQ. ID No. 6) for the porcine lactoferrin protein.

Fig. 16 is a Western blot showing hLF expression in *Saccharomyces Cervisiae.*

Fig. 17 is a schematic of the plasmid used for expression of the cDNA (SEQ. ID No. 1) in *Aspergillis Nidulans.*

Fig. 18 shows restriction enzyme cleavage sites for the human cDNA sequence.

Fig. 19 shows restriction enzyme cleavage sites for the bovine cDNA sequence.

Fig. 20 shows restriction enzyme cleavage sites for the porcine cDNA sequence.

DETAILED DESCRIPTION OF THE INVENTION

DEFINITIONS

For the purposes of the present application, the term "transferrin family" means a family of iron transferring proteins including serum transferrin, ovotransferrin and lactoferrin. These proteins are all structurally related.

For the purposes of the present application, the term "vector(s)" means plasmid, cosmid, phage or other vehicle to allow insertion, propagation and expression of lactoferrin cDNA.

For the purposes of the present application, the term "host(s)" means any cell that will allow lactoferrin expression.

For the purposes of the present application, the term "promotor(s)" means regulatory DNA sequences that controls transcription of the lactoferrin cDNA.

For the purposes of the present application, the term "multiple cloning cassette" means a DNA fragment containing restriction enzyme cleavage sites for a variety of enzymes allowing insertion of a variety of cDNAs.

For the purposes of the present application, the term "transformation" means incorporation permitting expression of heterologous DNA sequences by a cell.

For the purposes of the present application, the term "iron binding capacity" means ability to bind Fe. Fully functional human lactoferrin can bind two atoms of iron per molecule of LF.

For the purposes of the present application, the term "biological activity/biological active" means biological activity of lactoferrin as measured by its ability to bind iron, or kill microorganisms, or retard the growth of microorganisms, or to function as an iron transfer protein.

For the purposes of the present application, the term "substitution analog" referring to a DNA sequence means a DNA sequence in which one or more codons specifying one or more amino acids of lactoferrin or a lactoferrin polypeptide are replaced by alternate codons that specify the same amino acid sequence with a different DNA sequence. Where "substitution analog" refers to a protein or polypeptide it means the substitution of a small number, generally five or less, commonly 3 or 4, and more often 1 or 2 amino acids as are known to occur in allelic variation in human and other mammalian proteins wherein the biological activity of the protein is maintained. For example, hLF isolated from milk has been reported to differ from the hLF of SEQ. ID No. 2 at two amino acid residues.

The confirmation of the cDNA sequence and the deduced amino acid have been proven by multiple confirmation procedures.

These are:

1. Multiple sequence analyses.
2. Comparison of the amino acid sequence deduced from the cDNA with that of hLF generated by conventional amino acid sequencing of hLF isolated from milk. The unique cDNA sequence which encodes the human lactoferrin

protein has a variety of applications as known and indicated in the literature.

3. Transcription and translation of hLF protein from the cDNA with positive identification using an anti-hLF antibody.

The cDNA sequence of the present invention can be used to prepare recombinant human lactoferrin, thus making available a source of protein for therapeutic and nutritional applications. The confirmed cDNA of this invention can be used in an appropriate cloning vehicle to replicate the cDNA sequence. Also, the cDNA can be incorporated into a vector system for human lactoferrin expression. Other lactoferrin DNA sequences can be substituted for the human lactoferrin cDNA sequence to provide bovine, porcine, equine or other lactoferrins. Partial cDNA sequences can also be employed to give desired lactoferrin derived polypeptides. The expression systems of the invention can be used to provide lactoferrin derived polypeptides that are not available by enzymatic digestion of naturally occurring lactoferrin. The invention further provides an expression system for producing lactoferrin and lactoferrin related polypeptides in mammalian cell lines, other eucaryotic cells including yeast and fungal cells and procaryotic cells. The invention allows for the production of lactoferrin free of lactoperoxidase, lysozyme, or other proteins that are contaminants of lactoferrin isolated from milk or other natural products. This invention is not limited to any particular uses of the human cDNA sequence or production of lactoferrin of other species from the appropriate DNA sequences.

The recombinant LF being a protein derived by recombinant techniques can be used in a variety of applications. The human gene can be transferred to mammalian systems such as cows and other agriculturally important animals and expressed in milk. The incorporation of a human lactoferrin gene and expression in the milk of animals can combat an iron deficiency typical in piglets. The inclusion of the human lactoferrin gene with expression should improve an animal's disease resistance to bacterial and viral infection. The tissue specific expression of human lactoferrin in mammary glands, for instance, would impart

the bacteriocidal and virucidal benefit of the expressed gene to young feeding on the milk and would provide a production means for the secreted protein for therapeutic use.

The gene can be placed in the appropriate cloning vector for the production of LF. The LF produced by recombinant methods can be used in a variety of products including human or animal foods, as therapeutic additives to enhance iron transport and delivery, and for the virucidal and bacteriocidal qualities, as additives for eyedrops, contact lens and other eye care solutions, topical skin care products, eardrops, mouthwashes, chewing gum and toothpaste. The recombinant LF would provide a safe, naturally occurring product which can be topically applied as well as ingested safely. The bactericidal lactoferrin polypeptides are useful as preservatives in the above listed products, and as therapeutic anti-infection agents. The iron binding polypeptides are useful as iron carrier proteins for nutritional and therapeutic uses, and as bacteriostats and bactericides, especially in products of the types listed above. Each protein may also be used as a nutrition supplement and as a source of amino acids.

The full-length cDNA encoding human lactoferrin has been isolated, and the analysis has been completed. The cDNA sequence has been confirmed as human lactoferrin cDNA by comparison of the deduced amino acid sequence with the published amino acid sequence of hLF. The expression of lactoferrin was observed in a eucaryotic expression system from the cDNA and a plasmid vector. The presence of lactoferrin was confirmed by standard Western immunoblot analysis using anti-human lactoferrin antibodies and relative molecular mass measurement.

Fig. 1 is a schematic of the lactoferrin cDNA. The sequence can generally be described as an initial 5' untranslated region, 17 nucleotides in length. The next portion is 57 nucleotides which codes for the 19 amino acid secretion signal peptide starting with methionine. The next sequence of the cDNA codes for the mature human lactoferrin protein of 692 amino acids followed by the 3' untranslated region of 208 nucleotides which ends the cDNA. The complete

sequence is 2,358 nucleotides in length. The hLF protein contains glycosylation sites. The hLF protein with secretion signal sequence has an expected molecular mass of 78,403 daltons and the mature hLF is 76,386 daltons without added carbohydrate from glycosylation.

Fig. 2 is the cDNA sequence (SEQ ID No. 1) with the deduced amino acids (SEQ ID No. 2) for the secretion signal peptide and the mature human lactoferrin protein. The numbers on Fig. 2 correspond to the nucleotides starting at the 5' end. There are binding sites for two iron atoms with four amino acids participating in the binding of each iron. The amino acids at positions Asp80, Tyr112, Tyr209, and His273 are required for coordination with one iron, and amino acids at positions Asp415, Tyr455, Tyr548, and His617 bind the other. There are two glycosylation sites at positions Asn157 and Asn498. The numbers refer to the deduced amino acid sequence. There are 25 amino acids per line of protein sequence (starting at nucleotide 18).

The nucleotide sequence analysis was performed on cDNA isolated from a human prostate cDNA library. The prostate cDNA library yielded a 2,140 bp cDNA which contained the complete 5' end including the untranslated portion and the signal sequence. The 3' end including the three amino acids at the carboxy terminal and the untranslated region were obtained as a 208 bp cDNA from both a monocyte cDNA library and human prostate cDNA library.

The data in Fig. 2 displays the full-length cDNA sequence of this invention. The complete sequence including the 5' untranslated region and signal peptide have not been reported. Further, the previously reported amino acid sequence varies from the deduced amino acid sequence for hLF of this invention. The following TABLE 1 is a summary of the differences of the amino acid sequence of the present invention and those reported by Metz-Boutigue *et al.*, *Eur. J. Biochem.*, vol. 145, pp. 659-76 (1984). For the purpose of this table, the numbering of the amino acids will be initiated with methionine at the start of the signal peptide sequence as amino acid #1.

TABLE 1
COMPARISON OF AMINO ACID SEQUENCES
HUMAN LACTOFERRIN

Amino Acid Deduced from cDNA of hLF	Change	Metz-Boutigue Sequence
# 30 Thr	Substitution	Ala
# 48 Arg	Substitution	Lys
# 141 Arg	Insertion	NONE
# 170 Ala	Insertion	NONE
# 204 Ser	Substitution	Leu
# 206 Gln	Substitution	Lys
# 209 Tyr	Substitution	Lys
# 386 Glu	Substitution	Gln
# 392 Ser	Substitution	Trp
# 410 Asp	Substitution	Asn
# 411-424	Deletion	13 Amino acids in protein sequence not in deduced amino acid sequence from cDNA
# 532 Gln	Substitution	Glu
# 695 Lys	Substitution	Arg

Fig. 3 is the expression of human lactoferrin protein from the complete hLF cDNA. In addition to using the entire cDNA sequence and deduced amino acid sequence, a polypeptide of less than the entire protein can be of value. For instance, the region between amino acids 74-275 contains an iron binding domain which may be used without the rest of the protein for biologically available iron or the bacteriostatic qualities.

The cDNA sequence has been confirmed to encode lactoferrin. The hLF cDNA was shown to encode lactoferrin by expression of the cDNA in a eucaryotic expression system and detection of the expressed lactoferrin protein by Western immunoblot analysis using specific lactoferrin antibodies.

Recombinant production of lactoferrin protein has been described below in its preferred embodiments. However, it is also produced in a number of other sources such as fungal sources such as *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, or *Pichia pastorsis*, or insect cells such as SF9, or bacterial cells such as *Escherichia coli, or Bacillus subtilis*.

In one embodiment of the present invention, biologically active recombinant lactoferrin protein is produced. This method comprises synthesizing sequences containing a selectable marker gene, a promotor, a transcription termination sequence and a linker sequence.

Subsequently, the sequences are cloned to form a plasmid and the plasmid is digested with a restriction endonuclease. A cDNA coding for lactoferrin is inserted into a restriction site and eucaryotic cells are then transformed with the plasmid expressing the lactoferrin cDNA.

The selectable marker gene useful in the method of the present invention may be any that permits isolation of cells transformed with a lactoferrin cDNA plasmid. Preferably, the selectable marker gene is selected from *pyr4, pyrG, argB, trpC* and *andS*.

The promotor useful in the present invention may be any that allows regulation of the transcription of the lactoferrin cDNA. Preferably, the

promotor is selected from the group of alcohol dehydrogenase, argB, α-amylase and glucoamylase genes.

The transcription termination sequence useful in the present method may be any that allows stabilization of the lactoferrin mRNA. Preferably, the transcription termination sequence is derived from the α-amylase, glucoamylase, alcohol dehydrogenase or *benA* genes.

The linker sequence useful in the present method may be any that contains a translation initiation codon, a secretory signal and a restriction enzyme cleavage site. Preferably, the linker element is derived from the α-amylase or glucoamylase genes.

The cells, preferably eucaryotic cells, useful in the present invention are any that allow for integration of a vector, preferably a plasmid comprising the lactoferrin cDNA and expression of the lactoferrin cDNA. Preferably, the eucaryotic cells are fungal cells or insect cells. Insect cells such as SF9 are useful in the method of the present invention. More preferably, the fungal cells are yeast cells or *Aspergillus*. Most preferably, the eucaryotic cells useful in the present invention are *Aspergillus* strains, such as *A. oryzae*, *A. niger*, *A. nidulans* and *A. awamori*.

The invention also comprises partial sequences of the cDNA of SEQ ID No. 1, 3 and 5 and substitution analogs thereof which code for biologically active polypeptides having homology with a portion of lactoferrin, especially those that are not available from enzyme digests of natural lactoferrins, the method of making polypeptides by use and expression of partial cDNA sequences, and the polypeptide products produced by the methods of this invention. The desired partial sequences can be produced by restriction enzyme cleavage, as for example at the cleavage sites indicated in Figures 18, 19 and 20. the partial sequences may also be synthesized or obtained by a combination of cleavage, ligation and synthesis, or by other methods known to those skilled in the art.

Recombinant production of lactoferrin protein and polypeptides has been described in its preferred embodiment. However, it is also produced in a number of other sources such as fungal sources such as *Saccharomyces cerevisiae, Kluyveromyces lactis,* or *Pichia pastorsis* or insect cells such as SF9, and lactoferrin polypeptides may also be produced in bacterial cells such as *Escherichia coli, or Bacillus subtilis.*

The following examples are given for the purposes of illustrating various embodiments of the present invention and are not meant to be limitations of the present invention in any form.

EXAMPLE 1

HUMAN LACTOFERRIN cDNA

The complete 2,358 bp hLF cDNA was ligated to the eucaryotic expression vector, p91023(B) at the EcoRI site downstream from the adenovirus major late promoter. This plasmid vector was provided by Genetics Institute (Cambridge, Massachusetts) and has been described in previous publications (Wong *et al.*, *Science* 288:810-815 (1985)). The hLF cDNA expression vector was transferred into COSM-6 monkey kidney cells using standard tissue culture transfection conditions (Wigler *et al.*, *Cell*, 16:777-785 (1979)). These COS cells do not normally express lactoferrin. Forty-eight hours after transfection, the cells were harvested and crude cell extracts were prepared. Positive identification of the human lactoferrin was made by standard Western immunoblot analysis of the proteins expressed in the cell extracts, as well as those secreted into the cell growth medium using a commercially available antibody directed against human lactoferrin (Sigma). Proteins which bound to the anti-lactoferrin antibody were detected using radio-iodine labelled Protein A which reacts with the antibody. The immunoblots were autoradiographed to identify the human lactoferrin protein. Fig. 3 is an autoradiographic film showing the human lactoferrin expressed in four cell extracts prepared from tissue culture cells which were transfected with the lactoferrin cDNA expression vector (lanes 5 to 8). Lanes 5 to 8 show that the transfected cells all contain human lactoferrin (marked with an arrow) which is

immunoreactive with the anti-lactoferrin antibody and is the same molecular weight as human lactoferrin (M_r = 78,403 daltons). The control cells which were not transfected with the cDNA did not contain lactoferrin (lanes 3 and 4). Analysis of the growth medium showed that human lactoferrin was also secreted into the medium from transfected cells (lane 2) but not from control cells (lane 1).

The cDNA encodes a recombinant human lactoferrin protein which is similar to human lactoferrin protein isolated from milk as determined by molecular size comparisons and immunoreactivity with anti-human lactoferrin. Furthermore, the secretion signal peptide sequence is functional since the human lactoferrin is secreted into the growth medium of tissue culture cells which express the cDNA.

Fig. 4 is a schematic representation of the human lactoferrin protein precipitated after *in vitro* transcription and translation of the human lactoferrin cDNA. The 2140 bp cDNA was from the human prostate cDNA library and included the 5' untranslated region and the rest of the base pairs correlative to the cDNA sequence of Fig. 2 omitting the last 208 bp at the 3' terminus. The 2140 bp cDNA was ligated to the EcoRI site of the plasmid vector $pGEM_4$ (commercially available from Promega Biotech., Madison, WI 53711-5305) downstream from the SP_6 promoter. The plasmid construct was linearized at the 3' end of the hLF cDNA using the restriction enzyme Hinc II or Xba I. The linear DNA template was then transcribed *in vitro* using purified SP_6 RNA polymerase in the presence of ribonucleotides as described in the manufacturers protocol (Promega Corporation 1988/1989 Catalogue and Applications Guide). The resultant mRNA was translated using 100ng mRNA template and micrococcal nuclease treated rabbit reticulocyte lysate (as described by Promega) in the presence of 75uCi ^{35}S methionine (800 ci/mmol, Amersham). *In vitro* synthesized lactoferrin was immunoprecipitated by incubating 100ul aliquots of translation reaction with 10ug of rabbit anti-human lactoferrin IgG (Sigma Chemical Company, St. Louis, MO 63178) for 2 hours at 4°C in 50mM Tris, pH7.5/0.15M

NaCl/0.05% Tween-20 (1P buffer). The reaction volume was 200ul. Immunoreactive lactoferrin was precipitated after incubation for 1 hour with 50ug of Protein A sepharose (Pharmacia, Upsalla, Sweden). Immunoprecipitation was carried out by centrifugation for 5 minutes at 10,000g and the precipitate was washed 5 times with 4 volumes of 1P buffer. Total translation products and immunoprecipitates were then subjected to electrophoresis in denaturing 7.5% polyacrylamide gels. After fixing in 50% methanol, the gels were incubated in En^{3}Hance (NEN, DuPont, Wilmington, DE 19801) for 1 hour and washed with distilled H_2O. The gel was then dried under vacuum and exposed to Kodak X-OMAT XAR film at -70°C.

Lane 1 shows ^{14}C protein molecular weight markers used to estimate the size of the translated proteins. Lane 2 is a negative control which shows that no ^{35}S labelled proteins are translated in this system when no mRNA is added to the translation mix. Lanes 3 and 4 show the total translation products obtained when lactoferrin mRNA is added after preparation from two separate DNA templates. The major protein band (marked with an arrow) is human lactoferrin. This is the only band detected when the translation products are immunoprecipitated with anti-human lactoferrin before applying the protein to the gel (lane 6). The measurement of molecular mass by SDS-PAGE does not correspond to exact molecular weight due to secondary protein structure. However, the values are shifted in a correlative manner in comparison to the control. Analysis of the size of the translated lactoferrin is shown in Fig. 4. The protein migrated at the expected molecular mass of human lactoferrin (about 78Kd). The major bands in lanes 3 and 4 which migrate higher than the 68Kd marker band in the control lane correspond to expected molecular mass of hLF protein on SDS-PAGE.

EXAMPLE 2

FUNGAL STRAINS AND TRANSFORMATION

The *pyrG* mutant strain used in these studies was derived from *A. oryzae* (A07 11488). The *pyrG* gene from *A. oryzae* was mutated with 4-nitroquinoline-1-oxide. The *Aspergillus* transformation was carried out by, a modification of the procedure of Osmani, *et al.*, *J. Cell. Biol.* **104**:1495-1504 (1987). Conidia ($1X10^6$/ml) were inoculated into 50 ml of YG medium (0.5% yeast extract 2% glucose) containing 5 mM uracil and 10 mM uridine. Growth was at 32°C for 14-16 hours until a germ tube was visible. The germinated conidia were harvested by centrifugation and resuspended in 40 ml of lytic mix containing 0.4 M ammonium sulphate, 50 mM potassium citrate (pH 6.0), 0.5% yeast extract, 0.12 g novozyme, 0.1g Driselase, 100 μl β-glucuronidase, 0.5% sucrose and 10 mM $MgSO_4$. Protoplasting was for 2-3 hours at 32°C and 150 rpm. Following protoplasting, filtration using sterile miracloth was necessary to remove any undigested mycelia. The protoplasts were harvested by centrifugation and washed twice with 10 ml of 0.4 M ammonium sulphate, 1% sucrose and 50 mM potassium citrate (pH 6.0) at 4°C, resuspended in 1 ml of 0.6 M KCl; 50 mM CaCl; 10 mM Tris-HC1 (pH 7.5) and placed on ice. The transformation was performed immediately following the protoplast preparation. Aliquots (100 μl) of the protoplast were added to 3 μg of DNA and 50 μl of 40% polyethylene glycol (PEG) 6000, 50 mM $CaCl_2$, 0.6 M KCl and 10 mM Tris-HC1, (pH 7.5). The samples were incubated on ice for fifteen minutes after which an additional 1 ml of the PEG solution was added and incubation at room temperature was continued for thirty minutes. Aliquots of this mixture were plated in 3 mls of 0.7% minimal media, supplemented with 0.4% ammonium sulphate onto plates containing the same but solidified with 2% agar. All subsequent growth was at 32°C.

EXAMPLE 3

PLASMID CONSTRUCTION

A schematic representation of the expression plasmid is shown in Fig. 5. The complete cDNA encoding human LF was repaired using the Klenow fragment of DNA polymerase I and subcloned into Acc I digested and repaired pGEM4 to generate pGEMhLFc. In order to remove the LF signal sequence and generate a 5' end in frame with the α-amylase sequences, a 252 base pair lactoferrin fragment (nt 69-321) containing Hind II/Acc I ends was obtained by polymerase chain reaction (PCR) amplification of pGEMhLFc plasmid DNA. The oligo primers used were as follows: the 5' end oligonucleotide as shown in SEQ. ID. No. 7:

(CTGGGTCGACGTAGGAGAAGGAGTGTTCAGTGGTGC)

and the 3' end oligonucleotide as shown in SEQ. ID. No. 8:

(GCCGTAGACTTCCGCCGCTACAGG).

This PCR fragment was digested with Hind II and Acc I and was subcloned into Hind II/Acc I digested pGEMhLFc generating pGEMhLF. A 681 base pair α-amylase fragment with Asp718/Pvu II ends encoding the promotor, signal sequence and the alanine residue from the start of the mature α-amylase II gene, was obtained by PCR amplification of *A. oryzae* genomic DNA. The oligo primers were as follows: the 5' end oligonucleotide as shown in SEQ. ID. No. 9:

(GAGGTACCGAATTCATGGTGTTTTGATCATTTTAAATTTTTATAT)

and the 3' end oligonucleotide as shown in SEQ. ID. No. 10:

(AGCAGCTGCAGCCAAAGCAGGTGCCGCGACCTGAAGGCCGTAC AG).

The amplified DNA was digested with Asp718 and Pvu II and subcloned into Asp718/Hind II digested pGEMhLF. The resulting plasmid (pGEMAhLF) was digested with EcoR I and the resulting 2.8 kb α--amylase-lactoferrin fragment was subcloned into a unique EcoR I site in pAL3 according to the method of generating pAhLF*. Synthetic oligonucleotides were used to provide the last five carboxy terminal codons of lactoferrin (nt 2138 - 2153) missing in pAhLF* and also to provide the first 180 bp of 3' untranslated sequences from the *A. niger* glucoamylase gene. The resulting plasmid (pAhLFG) was used to transform the *A. oryzae* pyrG mutant strain.

With reference to FIGURE 5, *Aspergillus oryzae* expression plasmid, pAhLFG contains 681 bp of 5'-flanking sequence of the *A. oryzae* AMY II gene which includes the signal sequence and first codon of mature α-amylase. The cDNA coding for mature human lactoferrin is subcloned in frame downstream from these sequences allowing recombinant protein production by the addition of starch to the growth medium. The *Aspergillus niger* glucoamylase 3' untranslated region provides the transcription terminator and polyadenylation signals. The plasmid also contains the *Neurospora crassa pyr4* selectable marker and an ampicillin resistance gene.

The plasmid construct (pAhLFG) used for expression of human LF contains a 681 bp fragment that encodes the promotor and secretory signal peptide of the *A. oryzae* α-amylase II gene (AMY II). The signal sequence also contains the codon for alanine from the start of the α-amylase mature protein generating the signal sequence cleavage site (Leu Ala) recognizable by an endogenase α-amylase peptidase. A human lactoferrin cDNA fragment encoding the mature protein was subcloned in frame immediately downstream from the AMY II sequences, placing it under the control of this highly efficient starch inducible promoter. In order to stabilize the transcribed human LF mRNA, a 180 bp fragment encoding the 3' untranslated region of the glucoamylase gene from

Aspergillus niger was ligated into a unique BamH I site in the multiple cloning cassette, immediately downstream of the human LF cDNA providing the transcription terminator and polyadenylation signals. The plasmid also contains the *Neurospora crassa* pyr4 selectable marker which complements a pyrG auxotrophic mutation of *A. oryzae* and allows for selection of spores that have been transformed with the plasmid by growth in the absence of uridine.

EXAMPLE 4

GENOMIC DNA MANIPULATION

A. oryzae DNA was isolated from 200 mg of lyophilized mycelia as described by Rasmussen, *et al.*, *J. Biol. Chem.*, **265**:13767-13775 (1990). The DNA was digested with EcoR I, size fractionated on a 0.8% agarose gel and transferred to nitrocellulose. Prehybridization and hybridization of the nitrocellulose filter for Southern analysis were performed in 6XSSC, 0.1% SDS and 0.5% dried milk at 65°C for 16 hours. Hybridization solution contained 1 x 10^7 cpm ^{32}P-labelled lactoferrin cDNA probe (2.1 Kb). The filter was washed in 2XSSC, 0.5% SDS at room temperature for 30 minutes followed by two washes in 0.5X SSC, 0.5% SDS at 68°C for 30 minutes. The filter was dried, exposed at -70°C for two hours and developed by autoradiography.

With reference to FIGURE **6**, Southern blot analysis was performed on transformed *Aspergillus oryzae* strains. Genomic DNA from individual transformants and control AO7 were hybridized with a radiolabelled hLF cDNA probe (2.1 kb). The arrow points to a radiolabelled fragment (2.8 kb) generated upon EcoR I digestion of the expression plasmid which is present in all the transformants (#1-9) but is absent in control untransformed AO7. Molecular weights of bacteriophage lambda Hind III fragments are indicated at the left.

EXAMPLE 6

NORTHERN ANALYSIS

RNA was isolated from lyophilized mycelia (200 mg) using commercially available RNazol B (Biotecx Laboratories, INC, Houston,TX) according to the manufacturers instructions. Total RNA (20μg) was electrophoresed in a 0.8% agarose gel containing 2.2 M formaldehyde. The RNA was transferred to nitrocellulose and hybridized with either a 2.1 kb lactoferrin cDNA or a 1.8 kb genomic α-amylase fragment corresponding to the coding region of the α-amylase II gene. The probes were ^{32}P-labelled by nick translation (specific activity 2 X 10^8 cpm/ug). Hybridization was carried out 2 X SSC, .05% dried milk at 65°C over an ice with 2 x 10^6 cpm probe/ml.

Washes were identical to those employed in the Southern analysis. The filters were dried, exposed at -70° C for two hours and developed by autoradiography. RNA dot blots were performed using nitrocellulose membrane and the manifold dot blot system. Hybridization and washing conditions were as described above for Southern analysis. Radioactivity was quantitated using the betagon blot analyzer.

With reference to FIGURE 7, RNA analysis of transformant versus control AO7 was performed. In Panel A, Northern analysis of RNA (20 μg) from control AO7 and transformant #1 were hybridized with radiolabelled human LF cDNA. Human LF mRNA (2.3 kb) was detected in the transformant #1 but not in the control untransformed AO7. The positions of the 28S and 18S rRNA bands are indicated on the left. In Panel B, Dot blots of RNA (5 and 10 μg) from control AO7 versus transformant #1 using a radiolabelled α-amylase genomic DNA probe. In Panel C, Dot blots of RNA (5 and 10 μg) from control A07 and transformant #1 using radiolabelled human LF cDNA probe as illustrated.

Northern analysis was performed to determine if lactoferrin mRNA was transcribed correctly and efficiently in *A. oryzae* under the regulatory control elements of the expression plasmid. Spores (1x10^6/ml) from transformant #1 and from control untransformed spores were inoculated into fungal medium containing

1.5% glucose as carbon source and grown at 30°C for 48 hours in small shake flask cultures. The cultures were washed and reinoculated into fungal medium containing 3% starch to induce transcription of the human LF mRNA. After 24 hours, the cells were harvested and RNA was isolated. Total RNA (20 μg) was size fractionated on a 1.0% agarose gel containing 2.2 M formaldehyde and blotted on nitrocellulose.

Human lactoferrin mRNA was detected using ^{32}P labelled human LF cDNA (2.0 kb) probe. Hybridization with human LF radiolabelled cDNA probe detected a specific radiolabelled band at the correct size for lactoferrin mRNA (2.3kb) in the transformant but not in the control untransformed strain (Fig. 7A). Quantitation of mRNA levels by dot assay showed comparable levels of expression of endogenous α-amylase rRNA between control AO7 and transformant #1 (Fig. 7B). In addition, similar levels of expression of α-amylase and human LF mRNA were seen in transformant #1 (Fig. 7B and 7C).

EXAMPLE 6

PURIFICATION OF RECOMBINANT HUMAN LF

LF was purified from the growth medium using CM Sephadex C50 essentially as described by Stowell, et al., *Biochem J.*, **276**:349-59 (1991). The column was pre-equilibrated with 500 ml of 0.025 M Tris HCl, pH 7.50 1M NaCl. The pH of the culture medium was adjusted to pH 7.4 before applying to the pre-equilibrated column. The column was washed with 500 ml of equilibration buffer and followed by a linear salt gradient from 0.1 to 1.1 M NaCl. Fractions (7 ml total) were assayed for lactoferrin content and purity using SDS/PAGE and silver staining. Fractions containing LF were dialyzed against 0.025 M Tris HCl, pH 7.5/0.1M NaCl and lyophilized.

EXAMPLE 7

QUANTITATION OF HUMAN LF

Recombinant lactoferrin was quantitated using an ELISA assay essentially as described by Vilja *et al.*, *J. Immunol. Methods*, 76:73-83 (1985). A sensitivity of 5 ng of lactoferrin was obtained using the non-competitive Avidin-biotin assay. Human LF isolated from breast milk (Sigma) was used as standard. Biotinylated human lactoferrin IgG was obtained from Jackson Immunoresearch laboratories, West Grove, PA.

EXAMPLE 8

N-TERMINAL SEQUENCING

Five μg of purified recombinant human LF was resolved on an SDS-polyacrylamide gel and transferred to Problott, a polyvinylidene difluride-type membrane, following manufacturers instructions (Applied Biosystems). Human LF was detected with Comassie Brilliant Blue staining and destained. This human LF band was excised, washed thoroughly with distilled H_2O and air-dried. The N-terminal amino acid sequence of the first ten amino acids of human LF was determined by the automated Edman degradation procedure using an applied Biosystems Pulsed-liquid phase sequencer (Model 477A).

With reference to FIGURE 8, panel A illustrates a Silver stained SDS-polyacrylamide gel analysis of recombinant human LF secretion and purification. Lane 1 contains breast milk human LF standard (500 ng). Lanes 2 and 3 contain samples of the growth medium (40 μg) from induced control AO7 and transformant #1 respectively. Lanes 4-8 contain 100 μl aliquots of eluted fractions (#25, 30, 35, 40, and 45 respectively) collected from the CM-sephadex purification of recombinant LF from the growth medium of transformant #1. The position of the molecular weight markers (BioRad Richmond, CA) are indicated on the left. Sizes are given in kilo Daltons. Panel B illustrates a Western immunoblot analysis of duplicate samples as described in panel A using a specific polyclonal antibody directed against human LF with detection with ^{125}I-protein A.

Panel C illustrates #6 N-terminal amino acid sequence of recombinant human LF. Recombinant human LF was sequenced from the N-terminus through 10 residues and is identical to breast milk human LF with the exception of the additional alanine generated in our construction to provide the α-amylase signal sequence cleavage site.

EXAMPLE 9

DEGLYCOSYLATION

Deglycosylation was performed using N-glycosidase F (Boehringer Mannheim). *A. oryzae* growth medium containing 0.5 μg lactoferrin was denatured for 3 minutes at 100° C in the presence of 0.01% SDS. Standard LF from human milk was treated similarly. The samples were subsequently placed on ice for five minutes. N-glycosidase F reactions were conducted in 0.4 M sodium phosphate, (pH 6.8); 0.08% Triton; 0.1% β-mercaptoethanol and 1 unit of enzyme and incubated at 37°C for sixteen hours. PAGE and Western analysis was performed using an IgG specifically directed against human lactoferrin to detect an increase in mobility of digested samples.

With reference to FIGURE 9, recombinant human LF was characterized. Panel A illustrates the deglycosylation of lactoferrin. Western analysis of glycosylated and deglycosylated lactoferrin using a specific polyclonal antibody was directed against human lactoferrin with detection with ^{125}I-protein A. The first panel contains authentic breast milk human LF (500 ng) untreated (-) and treated (+) with N-glycosidase F. The second panel contains purified recombinant human LF (500 ng) untreated (-) and treated (+) with N-glycosidase F. The size of glycosylated human LF is indicated with the arrow. Panel B illustrates a functional analysis of recombinant lactoferrin with regard to iron-binding capacity. Panel A and B show the ^{59}Fe filter binding assay of duplicate samples of authentic breast milk human LF and purified recombinant human LF, respectively, at the concentrations indicated. The first lane in both panels contain BSA (5 μg) as a negative control.

Lactoferrin contains two N-acetyllactamine type glycans attached through N-glycosidic linkages. To determine if recombinant lactoferrin was glycosylated correctly, the protein was treated with N-glycosidase F, resolved on SDS-polyacrylamide electrophoresis, transferred to nitrocellulose and probed using a specific IgG directed against human lactoferrin (Fig. 11A). N-glycosidase F hydrolyses at the glycosylamine linkage generating a carbohydrate free peptide of smaller molecular weight. Comparison of recombinant LF with purified LF from human milk, illustrates that both proteins co-migrate upon digestion with N-glycosidase F suggesting that the recombinant protein has a glycosylation pattern similar to native LF.

Lactoferrin has a bilobal structure with each lobe having the capacity to bind tightly, but reversibly, one Fe^{3+} ion. The iron-binding properties of lactoferrin are crucial for its functional roles. To test if recombinant human LF expressed and secreted in *A. oryzae* has an iron binding capacity similar to authentic lactoferrin, an ^{59}Fe micro filter binding assay was developed. Purified human lactoferrin isolated from the growth medium of transformant #1 was dialyzed against 0.1M citric acid (pH 2.0) to generate apo-human LF. Native lactoferrin from human milk was treated similarly. Excess ^{59}Fe (0.2 mCi) was added to these samples in an equal volume of 1 M bicarbonate, followed by incubation at 37°C for 30 minutes. Samples were applied to nitrocellulose membrane and washed several times with bicarbonate. The filter was visualized by autoradiography and Fe-binding was quantitated using a betagon blot analyzer. As illustrated in Fig. 11B, both recombinant and native LF showed a similar level of iron binding at all concentrations tested. The results demonstrate that recombinant human LF is indistinguishable from native human LF in its capacity to bind iron.

With reference to FIGURE 2, the complete cDNA sequence for human lactoferrin protein is depicted. The cDNA coding for lactoferrin is used to create plasmids and transform eucaryotic cells and to produce the lactoferrin protein.

Strains of *Aspergillus* used in the present invention are auxotrophic mutants that contain a defective *pry4* gene that results in an inability to synthesis orotidine 5' phosphate (OMP) decarboxylase. The enzyme is required for uridine synthesis. The strain cannot grow on media lacking uridine. The plasmid contains a selectable marker, i.e., a sequence that encodes the gene for OMP decarboxylase. Uptake of the plasmid by the *Aspergillus* can therefore be selected for by growth on media lacking uridine. The *Aspergillus* is transformed by the plasmid such that it can grow on the uridine deficient media.

EXAMPLE 10

EXPRESSION OF THE 3' IRON-BINDING DOMAIN OF HUMAN LACTOFERRIN - *E. COLI*

The 3' iron-binding domain of human lactoferrin (hLF) was expressed in *Escherichia coli* using the bacterial expression plasmid, PT7-7 as described by Tabor, S. and Richardson, C., *Proc. Natl. Acad. Sci. U.S.A.*, **82**:1074-1078 (1985). pGEMhLFc, containing the cDNA for the complete hLF cDNA (Ward, P.P., *et al. Gene.* **122**:219-223 (1992)), was digested with Sma I and Hind III to release a 1.5 kb fragment encoding the 3' iron-binding domain of hLF. This 1.5 kb Sma I/Hind III fragment was subcloned in-frame into Sma I/Hind II digested PT7-7, under the control of the strong inducible T7 promoter, generating PT7-7hLF3'.

PT7-7hLF3' was transformed into a protease deficient strain of *E.coli* which had previously been transformed with pGP1-2 plasmid which contained the T7 polymerase under the control of the λpL promoter as described by Conneely, O.M., *et al. In: Hormone Action and Molecular Endocrinology*. 5-48 - 5-50 (1989)). The PT7-7 plasmid contained an ampicillin resistance gene while the pGP1-2 plasmid contained a kanamycin resistant gene allowing dual antibiotic resistance selection for transformants containing both plasmids. Transformants obtained were cultured overnight in LB broth containing ampicillin (50μg/ml) and kanamycin (50 μg/ml) at 30°C/250 rpm. Overnight cultures were

subcultured into LB (500 ml) containing ampicillin and kanamycin and grown at 30°C/250 rpm until an $O.D._{600}$nm of 0.5-0.6 was obtained. At 30°C the λ repressor bound to the λpL promoter, thus blocking T7 polymerase production. Induction of the recombinant protein was achieved by raising the temperature to 42°C for one hour to inactivate the λ repressor thus allowing T7 polymerase production. The temperature was lowered to 30°C for a further two hours, turning off λpL directed transcription and allowing the production of the recombinant protein as the T7 polymerase bound to the T7 promoter to specifically induce expression of the recombinant lactoferrin 3′ iron-binding domain.

Western Immunoblot analysis was performed to determine if the 3′ iron binding domain was expressed in the bacterial cells under the control of the T7 promoter and to monitor its purification. The cells were harvested at 5000*g* and resuspended in 15 ml of PBS (pH 7.4). Total cellular extracts were prepared by sonication for 1 minute on ice. The sonicate was centrifuged at 13,000*g* for 40 minutes at 4°C. The supernatant was removed and the pellet was resuspended in 50 ml of denaturation buffer (5M urea, 2% triton, 5mM EDTA, 0.01% Tween 20, 50 mM TrisCl, pH 7.5) and centrifuged at 48,000*g* for one hour. The supernatant containing the soluble fraction was recovered. Protein concentration was determined using the Bradford reagent according to manufacturers instructions (BioRad, Richmond, CA). Protein samples (40 μg) were resolved by SDS-PAGE and transferred to a nitrocellulose filter electrophorectically using the Western Immunoblot procedure. The filter was blocked with Tris-buffered saline (TBS, 0.05 M Tris/0.15 M NaCl, pH 7.5) containing 2% dried milk, and then incubated for 2 hours in the same with the addition of a specific polyclonal IgG (1 μg/ml) directed against hLF (Sigma, St. Louis, MO). The filter was washed (5 x 10 min) in TBS/0.05% Nonidet P-40 followed by incubation with 5 μCi of ^{125}I protein A in TBS/2% dried milk. The filter was washed (5 x 10 min) in TBS/0.05% Nonidet P-40, dried and exposed overnight in Kodak XAR5 film at -70°C. The film was developed by autoradiography.

The results of the Western analysis are shown in Figure 10. An immunoreactive band at the expected size (50 kDa) for the hLF 3′ iron-binding domain was evident in the cellular extract from induced cells and was absent in control uninduced cells (Figure 10, lanes 1 and 2). The hLF 3′ iron-binding domain associates with the cellular homogenate insoluble fraction (Figure 10, lane 3) and hence required a further solubilization step in a denaturation buffer to prepare the hLF in a soluble form (Figure 10, lane 4).

Analysis of a coomassie-stained SDS-PAGE gel also showed the presence of a 50 kDa protein in the cellular extract from induced cultures which was absent in control uninduced cultures (Figure 11, lanes 2 and 3). The recombinant protein was expressed at levels up to 10 mg/l and represented approximately 5% of the total cellular protein. The hLF 3′ iron-binding domain did not associate with the soluble homogenate fraction (Figure 11, lane 4) and hence required a further solubilization step in a denaturation buffer to prepare the hLF in a soluble form (Figure 11, lane 5). Purification and solubilization of the recombinant hLF 3′ iron-binding domain resulted in a 50% yield of recoverable protein and represented the major protein band in this fraction.

In summary, we have successfully produced recombinant hLF 3′ iron-binding domain in *E.coli* under the control of the strong inducible T7 promoter. The recombinant protein was expressed and purified in a soluble form from the cellular extracts at levels up to 5 mg/l.

EXAMPLE 12

EXPRESSION AND PURIFICATION OF AN N-TERMINAL LACTOFERRIN FRAGMENT (AA 1-52) IN *ESCHERICHIA COLI*.

An N-terminal human lactoferrin fragment (AA 1-52), encoding the bactericidal domain of hLF, reported by Bellamy *et al., supra,* was expressed and purified from *E. coli*. The bovine lactoferrin fragment also reported by Bellamy, *et al.* is produced by the same method illustrated here for the human fragment. This was achieved using the glutathione S-transferase (GST) Gene Fusion System

(Pharmacia, Piscataway, N.J.) where the lactoferrin fragment was expressed as a fusion protein with glutathione S-transferase [Smith, D. S., *et al.*, *Gene*, **67**:31-40 (1988)] and a protease cleavage site allowing production of the bactoricidal domain by cleavage from GST.

A 156 bp human lactoferrin fragment encoding AA 1-52, containing Sma I/BamH I ends was obtained by polymerase chain reaction (PCR) amplification of pGEMhLFc plasmid DNA [Ward, P. P., *et al.*, *Biotechnology*, **10**:784-789 (1992)]. The oligonucleotide primers used were as follows:

5' end oligonucleotide as shown in SEQ. ID. NO. 11

CTGCCCGGGCGTAGGAGAAGGAGTGTT

3' end oligonucleotide as shown in SEQ. ID. No. 12

CATGGATCCTGTTTTACGCAATGGCCTGGATACA

This PCR fragment was digested with Sma I and BamH I and repaired using the Klenow Fragment of DNA polymerase I. This fragment was subcloned into BamH I repaired pGEX-3X generating pGEX-3XLFN-1. This fused the lactoferrin cDNA fragment in frame, downstream from the glutathione S-transferase gene and under the control of the strong, inducible *tac* promoter. All PCR amplified products and construction junctions were sequenced using the commercially available Sequenase version 2.0 kit (United states Biochemical Corp, Cleveland, OH).

pGEX-3XLFN-1 was transformed into the bacterial strain, JM109. Transformants obtained were cultured overnight in LB (50 ml) containing ampicillin (50g/ml) at 37° C/250 rpm. Overnight cultures were subcultured into LB (500ml) containing ampicillin (50g/ml) and grown at 37° C/250 rpm until an $0D_{600}$nm of 0.6 - 0.8 was obtained. Isopropyl-D-thiogalactopyranoside (IPTG) was added to the culture medium at a concentration of 1 mM to turn on the *tac* promoter resulting in expression of the glutathione S-transferase/LFN-1 fusion protein. Growth under these conditions continued for 4 hours after which the cells were harvested at 5,000g and resuspended in 5 ml of MTPBS (150mM NaCl, 16mM Na_2HPO_4, 4mM NaH_2PO_4, 1% Triton X-100, pH 7.3). Total

cellular extracts were prepared by 3 X 1 minute freeze/thaw cycles followed by mild sonication for 2 X 1 minute. The sonicate was centrifuged at 13,000g for 20 minutes and the supernatant obtained was applied to a glutathione sepharose 4B column following manufacturer's instructions (Pharmacia, Piscataway, N.J.). The glutathione S-transferase/LFN-1 fusion protein was eluted from the column using 10 ml of elution buffer (10 mM glutathione, 50 mM Tris pH 8.0). Fractions of 1.5ml were collected and dialyzed overnight against 50 mM Tris, 15% glycerol pH 8.0.

Samples from the solubilized extracts and the purification fractions were analyzed by SDS/PAGE followed by silver-staining. The results of this analysis are shown in Figure 12. A band at the expected size (32 kDa) for the glutathione S-transferase/LFN-1 fusion protein was detected in the solubilized protein extracts from induced JM109 cultures transformed with pGEX-3X/LFN-1 and was absent in uninduced cultures (Figure 12A, lanes 2 and 3). This band migrates at a higher mobility than control induced JM109 cultures transformed with pGEX-3X alone (Figure 12A, lane 1). The fusion protein was successfully purified to homogeneity over a glutathione sepharose 4B column (Figure 12B, lanes 1 and 2). Protein concentration determination using the Bradford reagent (BioRad, Richmond, Ca) showed that the glutathione S-transferase/LFNI fusion protein was purified at levels up to 5mg/l. The GST fusion protein has a protease cleavage site for the protease Kex II between GST and the 52 amino acid protein.

In summary, a human lactoferrin fragment, encoding a bactericidal domain of this protein, has been successfully expressed as a fusion protein with glutathione S-transferase an *E. coli* expression system. This fusion protein was purified to homogeneity at levels up to 5mg/l. The bactericidal protein is obtained by cleavage with the protease Kex II to cleave the GST portion from the bactericidal domain.

EXAMPLE 13

EXPRESSION OF BOVINE AND PORCINE LACTOFERRIN IN *ASPERGILLUS ORYZAE*.

A universal *A. Oryzae* expression vector is constructed to allow in frame subcloning of any cloned cDNA of interest. This vector, pAG, is similar to the vector pAhLFG(+1) utilized for the expression of human lactoferrin in *A. Oryzae* above. A 680 bp α-amylase fragment encoding the promoter, signal sequence and the alanine residue from the start of the mature -amylase II gene, is obtained by polymerase chain reaction (PCR) amplification of pAhLFG(+1). The oligonucleotide primers are as follows:

5' end oligonucleotide, SEQ. ID. NO. 13

5'CGGAATTCATGGTGTTTTGATCATTTT

3' end oligonucleotide, SEQ. ID. NO. 14

5'TGGAATTCGATCGCGGATCCGCAATGCATGCAGCCAAAGCAGGTGCCG CGAC

The 5' end oligonucleotide encodes an EcoR I site and the 3' end oligonucleotide contains an Nsi I site, flanked by a BamH I site. This amplified DNA is digested with EcoR I and BamH I and subcloned into EcoR I/BamH I digested pAhLFG(+1) generating pAG. All PCR amplified products and construction junctions are sequenced using the commercially available Sequenase version 2.0 kit (United States Biochemical Corp., Cleveland, OH).

A schematic representation of this expression plasmid is outlined in Figure 13. Restriction enzyme digestion of this expression plasmid with Nsi I, followed by repair using DNA polymerase I allows subcloning of any cDNA of interest in frame with the α-amylase signal sequence and alanine residue from the start of the mature α-amylase II gene. 5' and 3' oligonucleotide primers are designed to contain Acc 1 ends, and used to obtain the full length cDNA encoding for mature porcine and bovine lactoferrin using polymerase chain reaction (PCR) amplification of their known DNA sequence. The PCR fragment thus obtained is digested with Acc I and repaired using the Klenow fragment of DNA

polymerase I for in frame subcloning into Nsi I blunt-ended pAG. The plasmids are then be transformed into the *pyr*G- strain of *A. Oryzae* to obtain expression and secretion of these cDNAs as previously described for human lactoferrin.

EXAMPLE 14

EXPRESSION OF HUMAN LACTOFERRIN IN *SACCHAROMYCES CEREVISIAE*

The complete human lactoferrin (hLF) cDNA was expressed in *Saccharomyces cerevisiae* using the yeast expression plasmid, YEP [McDonnell, D.P. *et al.*, *J. Steroid Biochem, Molec. Biol.*, 39:291-297 (1991)]. A 2.2kb fragment encoding the complete hLF cDNA SEQ. ID No. 1 was generated using the polymerase chain reaction. This fragment contained and *XhoI* restriction enzyme site at its 5' end and an *Asp718* restriction enzyme site at its 3' end. The 2.2kb fragment was subcloned, in frame, into *XhoI/Asp718* digested YEP to yield, YEPLFc.

Transcription of the hLF cDNA was under the control of the copper responsive yeast metallothionein promoter (CUP1). hLF was produced as a ubiquitin fusion protein. The fusion protein is short lived in the yeast cells and is processed to produce unfused protein upon folding.

YEPLFc was transformed into a protease deficient strain of *S.cerevisiae*, by standard techniques [Ito, H., *et al.*, *J. Bacteriol.*, 153:163-186 (1983).] This strain cannot grow unless the growth medium is supplemented with adenine, uracil and tryptophan. The YEP plasmid contains a tryptophan selectable marker, thus, transformants were selected by tryptophan auxotrophy.

Transformants obtained were cultured overnight in selective medium containing 2% glucose, 0.1% casamino acids, 0.67% yeast nitrogen base, 0.001% adenine and 0.002% uracil at 30°C/200 rpm. When the cells reached an OD_{600}nm of 1.0, 1 X 10^6 cells were inoculated into 10ml of the selective medium and 100μm $CuSO_4$ added. The cells were grown for 24 hours at 30°C/200 rpm.

The purpose of adding the $CuSO_4$ was to induce expression of the hLF cDNA from the copper inducible CUP1 promoter.

Western immunoblot analysis was performed to determine if hLF was expressed in the yeast cells under the control of the CUP1 promoter. The cells were harvested by centrifugation at 5000xg for 5 min. and resuspended in 1ml of Z buffer (120mM $Na_2HP0_47H_2O$, 40mM $NaH_2PO_4H_2O$, 10mM KCl, 1mM $MgSO_47H_2O$, 0.27% 2-mercaptoehanol, pH 7.0). Total cellular extracts were prepared by glass bead homogenization. This procedure involved mixing the yeast cells with an equal volume of glass beads (0.5mm, B.Braun Instruments) and vortexing for 5 x 1 min. The homogenate was centrifuged at 13,000g for 10 min. and the supernatant removed. The protein concentration was determined using the Bradford reagent in accordance with the manufacturer's instructions (BioRad, Richmond, CA). Protein samples (50μg) were resolved by SDS-PAGE and electrophoretically transferred, overnight, to a nitrocellulose filter using the western immunoblot procedure. The filter was blocked with tris-buffered saline (TBS = 0.05M Tris/0.15M NaCl, pH 7.5) containing 1% dried milk and then incubated overnight, in the same, with the addition of a specific rabbit polyclonal antibody (1μg/ml) directed against hLF (Signa, St. Louis, MO). The filter was washed in TBS/0.1% Tween 20 (5 x 5 min.) followed by incubation with horseradish peroxidase (Amersham, UK) for 1 hour. The filter was washed in TBS/0.3% Tween 20 (3 x 5 min.) and then TBS/0.1% Tween 20 (3 x 5 min.). The filter was then treated with luminol and enhancer (Amersham, UK) for 1 min., dried and exposed for 1 min. to X-ray film. The film was developed by autoradiography.

These data demonstrate successful production of recombinant hLF in *S.cerevisiae* under the control of the copper inducible (CUP1) promoter.

The results of the western analysis are shown in Figure 16. An immunoreactive band at the expected size (78 kDa) for hLF was evident in the cellular extract from transformed *S. Cerevisiae* cells. Fig. 16, lane 1.

EXAMPLE 15

EXPRESSION OF hLF IN *ASPERGILLIS NIDULANS.*

Construction of the *Aspergillis Nidulans* Expression Plasmid.

The plasmid used for expression of *hLF* cDNA is shown schematically in Fig. 17. The cDNA of SEQ. ID No. 1 as a 2.3-kb clone contained the secretory signal sequence and complete translation frame. The sequence of the entire cDNA was confirmed by dideoxy sequence analysis (Sequenase version 2.0, U.S. Biochemical, Cleveland, OH). The cDNA was repaired using the PolIk and subcloned into *Acc*I-digested and blunt-ended pGEM4. The plasmid, pGEMhLF, was digested with *Hind*III + *Asp*718 and repaired using PolIk. The resulting 2.3-kb *hLF* fragment was subcloned into a unique *Sma*I site located in the multiple cloning cassette of pAL3 downstream from the *alc*A promoter, Waring, R.B., *et al., Gene*, **79**, 119-130 (1989), generating pAL3hLF. The β-tubulin transcription terminator fragment was obtained by digesting the 3'-untranslated region of the *benA* gene (nt 2569-2665; May et al., 1987) with *XbaI* + *NheI* and subcloned into *XbaI*-digested pAL3hLF generating pAL3hLFT. This plasmid was used to transform *A. nidulans* strain GR5 (*pyrG89*; *wa3*; *pyroA4*)

The *A.nidulans* expression plasmid, pAL3hLFT, contains 300 bp of 55' -flanking sequence of the *A. nidulans alcA* gene containing all the regulatory elements necessary for controlled gene expression. To construct pALhLFT, a 2.3-kb *hLF* cDNA fragment containing 17 nucleotides of 5' - *UTR*, the complete *hLF* ORF encoding the secretory signal peptide and mature hLF, followed by 209 nt of 3' *UTR* was subcloned into a unique *Sma*I site in pAL3 downstream from the *alcA* promoter. A 96-bp terminator fragment from the *A. nidulans* β-tubulin-encoding (*benA*) gene was subcloned into a unique *Xba*I site downstream from the *hLF* cDNA sequence. The plasmid also contains an Ap^R maker and the *N. crassa* pyr4 selectable marker (Waring et al., *supra,* 1989).

Transformation and Southern analysis

Transformation was carried out as described by May et al., *J. Cell Beol.*, 109, 2267-2274 (1989). Protoplasts were transformed with 3 μg of the expression plasmid with an efficiency of 40 transformants/μg DNA. Transformats obtained were purified three times through conidial spores. Southern blot analysis was performed to confirm that transformants contained integrated plasmid with *hLF* cDNA. A *hLF*-specific radiolabelled band was detected at the expected size (2.3 kb) in lanes 1-10 but not in DNA from control spores. These results demonstrate that *hLF* cDNA was integrated into the genome of all *A. nidulans* transformants tested and varied randomly from one copy (transformants Nos. 3, 6 and 10) to 20 copies (No. 5) per cell. The site of integration of the plasmid into the *A. nidulans* genome is random due to the absence of homologous sequences to target the vector into a particular site.

Southern blot analysis was conducted of transformed *A. nidulans*. Genomic DNA was isolated from ten individual *A. nidulans* (GR5) transformats and untransformed spores as described by Rasmussen, C.D.*et al., J. Biol. Chem.*, **265**, 13767-13775 (1990). The DNA (1 μg) was digested with *Eco*RI, size fractionated on a 0.8% agarose gel and transferred to a nitrocellulose filter and hybridized with a radiolabelled *hLF* cDNA probe (2.1-kb). A sample (20 ng) of *hLF* cDNA was used as a positive control (hLF cDNA). Prehybridization and hybridization of the filter was performed in 6 x SSC/0.1% SDS/0.5% dried milk at 65°C for 16h. The hybridization solution contained 200 ng of ^{32}P probe (2.1 kb; specific activity 4×10^8 cpm/μg of DNA). Filters were washed in

2 x SSC/0.5%SDS at 68°C for 30 min followed by 0.5 x SSC/0.5% SDS at 68°C for 30 min. The filter was dried and exposed to Kodak X-AR5 film at - 70°C for 30 min and developed by autoadiograpy. The autoradiography showed an intense 2.1 kb band for hLF.

Production of hLF in *Aspergillus nidulans*

Conidia (1 x 10^6/ml) were cultured in minimal media utilizing 100 mM Na acetate pH 6.5 as carbon source with or without addition of 1.2% ethanol to induce transcription of the *hLF* cDNA. GR5 was cultured as above except for the addition of 5 mM uridine and 10 mM uracil. Media and mycelia were harvested and separated using Miracloth (Calbiochem, San Diego, CA). Mycelia (200 mg) were freeze-dried and lyophilized overnight. Total cellular extracts were prepared by homogenization in a glass teflon homogenizer using 1 ml of phosphate-buffered saline (PBS; 137 mM NaCI/2.7 mM KCI/4.3 mM $Na_2HPO_47H_2O$/1.4 mM K_2HPO_4pH 7.4) in the presence of phenylmethylsulfonylfluorride (PMSF, 10 μg). The homogenate was centrifuged at 12000 x g for 30 min at 4°C and the supernatant containing the soluble fraction was recovered. The growth medium was concentrated by freeze drying and lyophilization and resuspended in 1/30 vol. in PBS pH 7.4. Protein concentration was determined using the Bradford reagent according to manufacturer's instructions (BioRad, Richmond, CA). Concentrated media samples containing 40 μg protein and soluble extracts (50 μg protein) were subjected to 0.1% SDS/7% PAGE, Laemmli, U.K., *Nature*, **227**, 680-685 (1970). Purified lactoferrin (250 ng, Sigma, St. Louis, MO) was used as standard (hLF std). The

resolved proteins were transferred to nitrocellulose filters electrophoretically using the Western blot procedure, Towbin, H., *et al., Proc. Natl. Acad. Sci. USA*, **76**, 4350-4354 (1979). Filters were blocked with Tris-buffered saline (TBS, 0.05 M Tris/0.15 M NaCl pH 7.5) containing 2% dried milk and then incubated by 2 h in the same with the addition of a 1 μg/ml of a specific polyclonal IgG directed against hLF (Sigma, St. Louis, MO). Filter washes (5 x 10 min) were in TBS/0.05% Nonidet P-40 followed by incubation with 1 μCi of [^{125}I] protein A in BS/2% dried milk. The filter was washed (5 x 10 min) with TBS/0.05% Nonidet P-40, dried and exposed overnight to Kodak XAR5 film at -70°C. The film was then developed by autoradiography. The autoradiographs demonstrate production of hLF. Western analysis was performed to determine if the *hLF* cDNA was expressed in the *A. nidulans* transformats under the control of the *alcA* promoter.

Conidia (1 x 10^6/ml) from transformat No. 5, which contained the highest number of copies of integrated *hLF* cDNAs, and from untransformed GR5 were inoculated into minimal medium utilizing glucose as the carbon source. After 18 h, the cultures were harvested, washed and reinoculated into minimal medium supplemented with 1.2% ethanol and grown for an additional 12 or 24 h before harvesting the cultures. Cell extracts and samples of the growth medium were resolved by SDS-PAGE, transferred to nitrocellulose and immunoblotted using a specific polyclonal IgG directed against hLF. An immunoreactive band indistinguishable from native hLF was evident in the cells and growth medium from transformat No. 5 after 12 and 24 h growth only after ethanol induction.

Cell extracts or growth medium obtained from untransformed GR5 did not contain an immunoreactive band even after addition of ethanol. These results demonstrate that *hLF* is expressed in transformed *A. nidulans* under the control of the *alcA* promoter.

Western analysis revealed hLF in the cells in all of the remaining transformants. In general there was a correlation between the plasmid copy number and the expression levels obtained. In the medium hLF was detected only with transformats containing multiple copies of integrated expressed plasmid (Nos. 1, 5, 7 and 10).

In order to monitor the levels of hLF produced in the system, a pilot fermentation of transformant No. 5 was carried out using the growth parameters described above. ELISA analysis, Vilja, P., *et al., J. Immunol. Methods*, **76**, 73-83 (1985), using a specific biotinylated IgG directed against hLF demonstrated that the total level of recombinant hLF produced was 5 μg/ml with approx. 30% (1.5-2.0 μg/ml) of this material secreted into the medium.

Iron binding analysis of hLF.

To test if recombinant lactoferrin synthesized and secreted in *A. nidulans* has an iron binding capacity similar to authentic human lactoferrin, samples of the growth medium of transformant No. 5 and untransformed GR5 spores were examined using an ^{59}Fe microfilter-binding assay to detect ^{59}Fe-bound lactoferrin. Iron-binding (^{59}Fe) is detected in the medium from transformant No. 5 but not in the medium from control untransformed GR5 spores. These results

indicate that hLF produced in *A. nidulans* is biologically active in its capacity to bind ^{59}Fe.

The data demonstrate the successful production of biologically active hLF in *A. nidulans*. The levels of hLF produced in *A. nidulans* were approx. 5 μg/ml with 30% of the hFL secreted into the growth medium. The secreted hLF was identical to native breast milk LF with regard to size and immunoreactivity. Furthermore, the hLF was capable of binding iron. Although hLF has been reported to contain anti-fungal properties, neither the re-hLF nor native hLF when added to the growth medium, retarded the growth of this strain of *A. nidulans*. The production of biologically active hLF in *A. nidulans* will facilitate testing of possible nutritional and therapeutic uses of this protein.

EXAMPLE 16.

PRODUCTION OF DNA SEQUENCE SUBSTITUTION ANALOGS.

Figure 18 shows the restriction enzyme cleavage sites in the SEQ I. D. No. 1 cDNA for cleavage by various endonucleases. Table 2 lists the alternative codons that code for the 20 common amino acids. DNA sequence substitution analogs that also code for human lactoferrin can be constructed by choosing alternate codons from Table 2 to alter the DNA Sequence between a pair of cleavage sites selected from Fig. 18. Alternative codons are assembled into a synthetic oligonucleotide by conventional methods and the synthetic oligo is substituted into the endonuclease treated DNA of Sequence ID. No. 1 by the methods described in "Molecular Cloning. A Laboratory Manual", 2d Edition, Cold Spring Harbor Laboratory Press (1989), to produce a substitution analog.

Other methods generally known to those skilled in the art can also be employed to obtain substitution analogs of DNA sequences. The alteration of the DNA by cleavage and codon substitution maybe repeated to substitute substantial portions of the original DNA sequence with alternative codons without altering the protein expressed by the DNA of Sequence ID. No. 1. The same methods can of course be used to make substitution analogs of the cDNA of SEQ ID No. 3 and 5. Alteration of a DNA sequence which produces no change in the protein expressed by the DNA sequence might, for example, be conducted to increase protein expression in a particular host cell by increasing the occurrence of codons that correspond to amino acid tRNAs found in higher concentration in the host cell. Such altered DNA sequences for substitution analogs can be easily produced by those of ordinary skill in the art following the method set out above, or other alternative techniques for altering the DNA sequence while obtaining the same protein on expression. Substitution analogs can be obtained by substitution of oligonucleotides at restriction cleavage sites as described above, or by other equivalent methods that change the codons while preserving the amino acid sequence of the expressed protein.

TABLE 2

AMINO ACID	CODONS
Phe	TTT TCC

AMINO ACID	CODONS
Leu	TTA TTG CTT CTC CTA CTG
Ile	ATT ATC ATA
Met	ATG
Val	GTT GTC GTA GTG
Ser	TCT TCC TCA TCG AGT AGC
Pro	CCT CCC CCA CCG
Thr	ACT ACC ACA ACG
Ala	GCT GCC GCA GCG
Tyr	TAT TAC

AMINO ACID	CODONS
Gly	GGT GGC GGA GGG
His	CAT CAC
Gln	CAA CAG
Asn	AT AAC
Lys	AAA AAG
Asp	GAT GAC
Glu	GAA GAG
Cys	TGT TGC
Trp	TGG
Arg	CGT CGC CGA CGG AGA AGG
TERMINATION SIGNALS	TAA TAG TGA

In conclusion, it is seen that the present invention and the embodiments disclosed herein are well adapted to carry out the objectives and obtain the end set forth in this application. Certain changes can be made in the method and apparatus without parting from the spirit and scopes of this invention. It is realized that changes are possible and that it is further intended that each

element or step presided in any of the filing claims is to be understood as to referring to all equivalent elements or steps for accomplishing the essentially the same results in substantially the same or equivalent manner. It is intended to cover the invention broadly in whatever form its principles may be utilized. The present invention, therefore, is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as others inherent therein.